Neurocomputing 131 (2014) 87–97

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Rapid multimodality registration based on MM-SURF

Dong Zhao, Yan Yang, Zhihang Ji, Xiaopeng Hu *

Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China

ARTICLE INFO

Article history:
Received 5 April 2013
Received in revised form
20 October 2013
Accepted 23 October 2013
Communicated by Jungong Han
Available online 22 November 2013

Keywords:
Multimodal-SURF
Multimodal images
MM-SURF
Multimodality registration
SURF

ABSTRACT

With a large number of registration algorithms proposed, image registration techniques have achieved rapid development. However, there still exist many deficiencies in multimodality registration where high speed and accuracy are difficult to simultaneously achieve for real-time processing. In order to solve these problems we propose a novel method named MM-SURF (Multimodal-SURF). Inheriting the advantages of the SURF, the method is able to generate a large number of robust keypoints. For each keypoint, the neighborhood gradient magnitude is utilized to compute its dominant orientation. Relying on the dominant orientation, a MM-SURF descriptor is constructed as the local features description of the keypoint. The geometric transformation matrix for multimodal image registration is obtained by matching the keypoints. The method makes full use of gray information of multimodal images and simultaneously inherits the good performance of the SURF. Experimental results indicate that the proposed method achieves higher accuracy and consumes less runtime than the other similar algorithms for multimodal image registrations, and also demonstrate its robustness and stability in the presence of image blurring, rotation, noise and luminance variations.

1. Introduction

Multimodal images, taken by different sensors at different times or from different viewpoints, can provide richer and more comprehensive information than monomodal images [1]. However, the increase of the number of image modalities also increases the level of difficulty for image registration to determine the geometric transform since the same part of an object can be represented by different intensities due to the different sensitivities of sensors in multimodal imaging [2]. To date, multimodality registration is still a challenging problem.

Multimodality registration techniques are generally classified into two categories, namely intensity-based registration and feature-based registration. In intensity-based registration techniques gray information is used to measure the similarity between images. The optimal transformation model is obtained by maximizing the similarity metrics computed by using methods including mutual information (MI) [3], cross correlation (CC) [8] and Fourier method [9]. The MI-based approaches [3–7], which use the statistical dependence of intensity values to measure the similarity between images, are commonly studied and have been widely applied in medical imaging. However, intensity-based techniques are usually susceptible to image noise due to their heavy dependence on the intensity relationship and ignorance of the structural information of images. Moreover, intensity-based methods are usually time-consuming owing to their huge search space of optimization.

In feature-based registration techniques, the geometric transform is determined by establishing the correspondence of image features such as edges and points. Once the correspondences of more than four features are obtained, the parameters of the transform can be computed [11]. In [10], the multispectral corner detectors are proposed to improve the performance of extracting interest points and a variant of the SIFT descriptor is employed for multimodality registration of near-infrared and visible-light images. In [11], straight lines derived from edge pixels are employed to estimate a global transformation, which is followed by a local adaptation process using a point-based transform. In [12], directed partial Hausdorff distance is applied to test transformation hypotheses of registering infrared and RGB images of buildings. In a video-based method proposed by Bilodeau et al., the trajectories of moving objects are obtained using background subtraction and tracking [13]. A registration criterion then is proposed for matched trajectory points by using RANSAC algorithm. For object motion detection, Cinque et al. propose a multimodality registration approach based on image segmentation and clustering [14]. Instead of matching image points, their method estimates the parameters of the projective transformation by matching edge and triangle clusters. In [15], line correspondence analysis is performed in the Hough parameter space of line segments for thermal and visual light image registration. Since the gray values of corresponding parts are usually not comparable

* Corresponding author.
E-mail address: xphu@dlut.edu.cn (X. Hu).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.10.037

over multimodal images, feature-based approaches have become the focus of research for multimodality registration. Compared with intensity-based methods, they are more robust to noise and less computationally expensive.

The key issues of feature-based methods are how to extract the corresponding features from two images and how to accurately match them. In multimodal images, the similar image structures mean that many of the same keypoints, at least those that do not depend on the texture, will tend to be detected [16]. Therefore, although many keypoints extracted from multimodal images are in different actual location, enough shared keypoints can still be detected for matching. In keypoint-extraction methods, SURF method can rapidly extract scale-invariant and robust keypoints [17]. It outperforms most of other image registration methods with respect to repeatability, distinctiveness, and robustness, and can be computed faster.

Gradient information has been used to design similarity measures for multimodality registration. In [18], an entropy-based similarity measure is suggested on the basis of a 3-D joint histogram incorporating intensity information and edge orientation information. In [19], a similarity measure based on directional-derivative energy images is proposed. In [20], an implicit similarity measure is proposed by summing radiant magnitudes within a mask. However, due to different contrasts between images, gradient magnitudes may not be correlated. To tackle this problem, Lee et al. present a concept of edginess to represent the level of confidence in the edge pixels [21]. When comparing different modal images, we observe a common phenomenon that the gradient directions of the corresponding points on the edge are same or opposite (see Fig. 1). The phenomenon has been used in several multimodality registration methods [10,16,18,22,23]. To verify this phenomenon, we have conducted several experiments. In Fig. 2, the locations of gradient reversals are highlighted in red. It can be found that most of gradient reversals happen around points containing a large amount of structural information that is critically needed for multimodality registration.

In this paper, we propose a novel SURF-based multimodality registration method named MM-SURF in order to reduce the computational cost and improve the accuracy. Due to the consistency of neighborhood structures in multimodal images, the directions around keypoints where gray values change severely are used as dominant orientations. To adapt for multimodality registration, the SURF descriptor is modified according to the gradient reversals. The proposed descriptor is more distinctive while maintaining its repeatability. Experiments indicate that the proposed MM-SURF method achieves good performance both in speed and accuracy for multimodality registration. Besides, MM-SURF is also proved to be still valid in the visible-light registration.

The rest of the paper is organized as follows. After a brief review of the SURF algorithm in Section 2, we present details of the Multimodal-SURF method (MM-SURF) in Section 3. The experimental results are shown and discussed in Section 4. We conclude this paper in Section 5.

2. Review of SURF

In 2006, Herbert Bay [17] proposed a speeded-up and robust image registration algorithm called SURF. The most important characteristic of SURF is fast, and simultaneously the algorithm also ensures the performances of repeatability, distinctiveness, and robustness. SURF employs the integral image, Haar wavelet response and approximate Hessian matrix to reduce computation time drastically and increase its robustness. SURF in general can be divided into four stages: detection and localization of keypoints,

Fig. 1. Example of gradient directions of multimodal images' corresponding points. The gradient directions of the red area are the same, while the gradient directions of the blue area are opposite. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

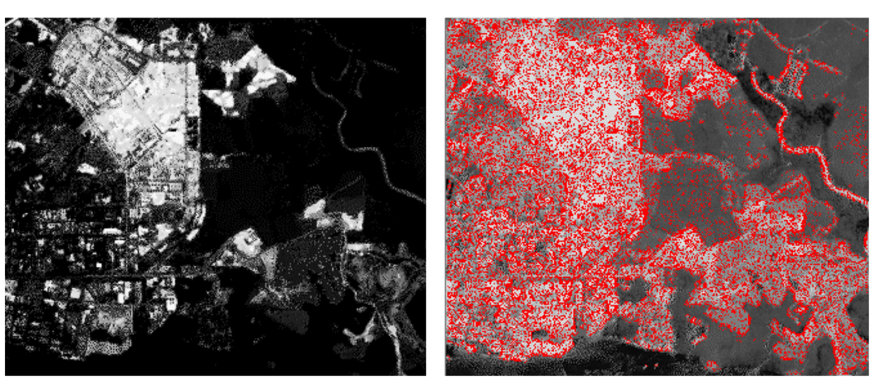

Fig. 2. Locations of gradient reversals in multispectral and visible spectrum images of a city. In the red positions where the gradient magnitudes are not zero, the gradient directions between two images are either the same or opposite. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)